



universidad
de león



FACULTAD DE CIENCIAS BIOLÓGICAS Y AMBIENTALES

**ANÁLISIS DE DATOS MEDIANTE TÉCNICAS DE
DEEP LEARNING APLICADAS A CASOS
MÉDICOS**

**DATA ANALYSIS THROUGH DEEP LEARNING
METHODS APPLIED TO MEDICAL CASES**

Autor: Miguel Cebrián Aldana

GRADO EN BIOTECNOLOGÍA

Junio de 2020



ÍNDICE

1	Introducción	1
2	Materiales y métodos	4
2.1	Bases de datos de mamografías	4
2.2	Regresión Lineal Múltiple	5
2.2.1	Concepto	5
2.2.2	$W_{xMaxima}$	6
2.3	Redes Neuronales	7
2.3.1	Perceptrón Simple	8
2.3.2	Perceptrón Multicapa	8
2.3.3	Adaline	9
2.3.4	Neuroph Studio	9
3	Resultados	11
4	Conclusiones	16
5	Referencias	17
	ANEXO 1. Ejemplo de funcionamiento de Neuroph Studio	19

Resumen

Actualmente, para la detección del cáncer de mama, la herramienta más eficaz es la mamografía, sin embargo, el elevado número de biopsias realizadas supone un problema, debido al gran porcentaje de biopsias realizadas innecesariamente que presentan resultados benignos. Para reducirlo, se han propuesto varios sistemas de diagnóstico basados en técnicas de deep learning a lo largo de los años, que proporcionan ayuda a los médicos para decidir si se realiza una biopsia de seno en caso de que aparezca una lesión sospechosa en la mamografía. En este trabajo se presentan dos métodos distintos, regresión lineal múltiple y redes neuronales artificiales, para predecir el resultado de la biopsia a través de distintos parámetros, tanto de la paciente, como de los que se aprecian en la mamografía. Se ha evaluado el rendimiento analítico de todos los métodos utilizados, mediante el error cometido, la especificidad y sensibilidad del método, y el análisis ROC de los modelos en el que se ha estimado una mejor predicción, que se trata del algoritmo perceptrón multicapa, en el que se ha conseguido una fiabilidad del 77% para la predicción por este método.

Palabras clave: redes neuronales, mamografía, cáncer de mama, análisis de datos, regresión lineal múltiple, diagnóstico.

Abstract

Currently, the most effective tool for the detection of breast cancer is mammography. However, the high number of biopsies performed is a problem, due to the large percentage of unnecessarily performed biopsies that present benign results. To minimize this problem, several diagnostic systems based on deep learning techniques have been proposed over the years, which provide assistance in the decision of doctors to perform a breast biopsy in case a suspicious lesion appears on the mammography. In this paper, two different methods, multiple linear regression and artificial neural networks, are presented to predict the result of the biopsy using different parameters, both of the patient and of those seen on mammography. The analytical performance of all the methods used has been evaluated, by means of the error committed, the specificity and sensitivity of the method, and the ROC analysis of the models in which a better prediction has been estimated, that is the multilayer perception method, in which a reliability of 77% for the prediction by this method has been achieved.

Key words: neural networks, mammography, breast cancer, data analysis, multiple linear regression, diagnosis.

1. INTRODUCCIÓN

Desde los primeros registros disponibles sobre la incidencia del cáncer de mama, que son de 1940, se registra una tendencia creciente en la incidencia de esta enfermedad, con incrementos anuales del 1-4%, tanto en los países desarrollados como en aquellos en vías de desarrollo (Merino Bonilla *et al.*, 2017).

El cáncer de mama es el cáncer más común, y además la principal causa de mortalidad por cáncer en mujeres en el mundo. Según la Organización Mundial de la Salud, la detección temprana sigue siendo la base para mejorar los resultados y supervivencia del cáncer de seno (Akram *et al.*, 2017). Relacionado con esto, las investigaciones científicas sobre cáncer de mama de las últimas décadas se han centrado en el diagnóstico precoz y su tratamiento, esto ha supuesto una revolución en el manejo de la enfermedad.

Entre lo más relevante, se encuentra la evolución tecnológica de los distintos métodos de detección, como la mamografía digital, los sistemas de lectura asistida por ordenador y la tomosíntesis. Además del desarrollo de un sistema estandarizado para la interpretación de las pruebas de imagen, la incorporación de técnicas quirúrgicas menos invasivas y los recientes avances en el campo de la genómica (Merino Bonilla *et al.*, 2017).

La idea de la detección temprana comenzó en Estados Unidos a principios del siglo XX, con campañas educativas masivas donde destacaba el mensaje de “do not delay” (no retrasar), sobre la búsqueda de ayuda médica para los síntomas del cáncer de mama (Aronowitz, 2007). La principal herramienta de detección temprana para cáncer de mama usada en los países occidentales, es la mamografía. El rango de edad de las pacientes a las que realizar estas pruebas de cribado varía según el país, pero suele ir de los 40 a los 74 años, y el intervalo recomendado entre dos pruebas es de 1 a 3 años (Løberg *et al.*, 2015).

Al igual que ocurre en todos los test de diagnóstico, la sensibilidad (capacidad una prueba para detectar una enfermedad en personas enfermas) y especificidad (capacidad de una prueba para dar negativos los casos de pacientes sanos), no son perfectas. El riesgo de experimentar un falso positivo durante una mamografía en Europa, para mujeres de 50 a 70 años es de aproximadamente del 20%, y además, el riesgo de realizar una biopsia por un falso positivo es del 3%, el riesgo de un falso positivo aún es mayor en Estados Unidos, donde es aproximadamente del 30% (Løberg *et al.*, 2015).

La mamografía es una técnica de detección temprana que puede identificar el cáncer de mama varios años antes de aparecer síntomas físicos, y es reconocido como el más efectivo, pero alrededor del 5% al 10% de los resultados, se interpretan como anormal o no concluyente. Para resolverlos se realizan otros exámenes como ecografía o biopsia de seno, que llevan a una interpretación final de si el tejido mamario es benigno o maligno, habiendo una probabilidad del 10% al 30% de que una biopsia de seno muestre un tumor maligno (Elter *et al.*, 2007). Esto, además de suponer un coste monetario, pueden influir negativamente en el bienestar psicológico de las pacientes y suponer una calidad de vida peor para ellas, e incluso se puede llegar a perder la confianza en el sistema de atención médica.

En los últimos años, se han propuesto sistemas de diagnóstico asistidos por ordenador, en los que se utiliza mayoritariamente el lenguaje de informe radiológico BI-RADS (Breast Imaging And Data System), que para reducir la variabilidad producida por el observador, se asigna una categoría de sospecha de carcinoma a cada hallazgo, lo que conlleva a una recomendación de actuación (Merino Bonilla *et al.*, 2017).

Mediante el lenguaje BI-RADS se deben analizar los siguientes parámetros en las mamografías de las pacientes:

- i. Densidades y masas.
- ii. Micro y/o macrocalcificaciones.
- iii. Distorsiones arquitectónicas.
- iv. Casos especiales que incluyen: ectasia ductal mamaria, ganglio intralinfático mamario, o densidad asimétrica focal.
- v. Hallazgos asociados: retracción de la piel o del pezón, engrosamiento de la piel, lesiones cutáneas, ganglios linfáticos axilares.

Al final del informe, se debe tener una conclusión, siguiendo las categorías de clasificación BI-RADS, que se pueden ver en la **figura 1** (Balleyguier *et al.*, 2007).

BIRADS 0	Incomplete, need for an additional imaging evaluation
BIRADS 1	Normal. Normal interval follow-up
BIRADS 2	Typically benign. Normal interval follow-up
BIRADS 3	Probably benign. A short interval follow-up is recommended: 4 months follow-up for masses and 6 months follow-up for microcalcifications
BIRADS 4	Suspicious abnormality: a biopsy should be considered
BIRADS 5	Highly suggestive of malignancy. Biopsy or surgery should be performed
BIRADS 6	Histologically proven malignancy. Imaging is performed for cancer staging or evaluation after chemotherapy

Figura 1. Categorías BI-RADS, donde se indica el significado de cada una de ellas, aplicadas a las imágenes mamográficas (Balleyguier *et al.*, 2007).

La aparición de nuevas herramientas de inteligencia artificial que pueden aprender y mejorar su rendimiento diagnóstico continuamente, han sido de gran interés para el diagnóstico basado en imágenes.

Los métodos de Machine Learning, utilizados en análisis de imágenes médicas, se dividen en dos grandes grupos, de aprendizaje no supervisado y de aprendizaje supervisado. Los primeros se basan en agrupar ejemplos similares, de acuerdo con alguna métrica de similitud predefinida anteriormente, y el objetivo es descubrir patrones en los datos que de otra forma serían difíciles de notar. Por otro lado, los métodos de aprendizaje no supervisado clasifican los datos sin conocer ninguna representación intermedia de estos, por lo que sólo funcionan bien si las características de entrada son muy predictivas. Un ejemplo de estos métodos, son los árboles de decisión (Reig *et al.*, 2019).

Otro ejemplo de estos métodos de aprendizaje no supervisado, son algunos de los modelos de redes neuronales artificiales, que están compuestos por unidades que reciben varios datos de entrada y producen una sola salida, estas unidades se denominan neuronas. Cada neurona, va a calcular la suma ponderada de las entradas, transformando el resultado de forma no lineal. Esta suma, se puede interpretar como la comparación del patrón de entradas con un patrón de pesos (Kriegeskorte y Golan, 2019).

2. MATERIALES Y MÉTODOS

A continuación, se especifican tanto la base de datos de casos de mamografías, como los programas informáticos y los métodos matemáticos utilizados para desarrollar y evaluar los sistemas CAD (computer-aided diagnosis), basados, en este caso, en la regresión lineal múltiple y en las redes neuronales artificiales.

2.1 Base de datos de mamografías

Para el desarrollo adecuado de estos métodos, es requerible obtener una base de datos de diferentes casos. Se ha elegido la base de datos de mamografías que se puede encontrar en el repositorio de datos para máquinas de aprendizaje de la Universidad de California en Irvine (UCI). Esta base de datos, contiene casos de pacientes identificados mediante mamografías digitales obtenidos en el Instituto de Radiología de la Universidad de Erlangen-Nuremberg, entre 2003 y 2006.

En la base de datos, hay un total de 961 casos y para cada uno de ellos, encontramos 5 atributos:

1. Evaluación BI-RADS, valores desde el 1 hasta el 5.
2. Edad de la paciente.
3. Forma de la masa: redonda = 1, ovalada = 2, lobular = 3, irregular = 4.
4. Margen de la masa: circunscrito = 1, microlobulado = 2, oscurecido = 3, mal definido = 4, espiculado = 5.
5. Densidad de la masa: densidad alta=1, isodensidad = 2, baja densidad = 3, contiene grasa =4.

Produciéndose como salida, el rango de severidad: benigno = 0 ó maligno = 1.

Todos estos parámetros se pueden ver en la **tabla 1**.

En esta base de datos, hay varios de estos valores que no se conocen, por lo que no se han utilizado los casos de esas pacientes. Por lo tanto, contamos con un total de 816 casos sin ruido, separados en dos conjuntos: el de entrenamiento, que contiene 670 casos, el 48,5% malignos y el 51,5% benignos, y el conjunto de validación, que constituye el 18% de la base de datos, y dentro de este, el 46,5% de los casos son malignos, y el 53,42% benignos.

Tabla 1. Muestras de los diferentes parámetros de la base de datos para cada paciente. (Dua, D. y Graff, C.,2019)

BI-RADS	Edad	Forma	Margen	Densidad	Severidad
5	67	3	5	3	1
5	58	4	5	3	1
4	28	1	1	3	0
5	57	1	5	3	1

2.2 Regresión Lineal Múltiple

En esta sección se va a desarrollar el método de regresión lineal múltiple, además de presentar el software utilizado para llevar a cabo este método.

2.2.1 Concepto

El modelo lineal simple se puede describir mediante la recta $Y = a_0 + a_1X$, donde a_0 es el coeficiente independiente y a_1 es la pendiente, de esta manera, si la pendiente es 0, no hay relación entre X e Y. Cuando hay más de una variable X, se puede extender geoméricamente a un hiperplano, en un espacio de dimensión $k+1$, donde k es el número de variables X. Por ejemplo, para un caso tridimensional, donde las variables independientes son X_1 y X_2 , el modelo matemático sería el siguiente:

$$Y = a_0 + a_1X_1 + a_2X_2.$$

(Bangdiwala, 2018)

Por lo tanto, para un análisis multivariable, el modelo matemático que se puede seguir es el siguiente:

$$Y = a_0 + \sum_{i=1}^{i=k} a_iX_i.$$

No obstante, por la complejidad de los datos utilizados, en las pruebas realizadas, estas variables X_i , van a ser funciones f_i , no lineales, que van a depender de 5 variables distintas (x,u,v,t,r), cada una de ellas se corresponde con un atributo de cada caso de la base de datos escogida anteriormente, por lo que el modelo va a ser el siguiente:

$$f(x, u, v, t, r) = a_0f_0(x, u, v, t, r) + \sum_{i=1}^{i=k} a_if_i(x, u, v, t, r),$$

donde x es el valor BI-RADS, u es la edad de la paciente, v es la forma de la masa, t es la forma del margen, y r es la densidad de la masa.

Para la realización de las pruebas de este trabajo, se han utilizado distintas funciones f_i , como se pueden ver en la **tabla 2**.

Tabla 2. Funciones utilizadas en el método de regresión lineal múltiple realizado con wxMaxima.

$f_i(x,u,v,t,r)$	1
$f_i(x,u,v,t,r)$	sen (x + u)
$f_i(x,u,v,t,r)$	v + t + r
$f_i(x,u,v,t,r)$	e ^ (v - x) + r
$f_i(x,u,v,t,r)$	(1 / u) + r
$f_i(x,u,v,t,r)$	v / r
$f_i(x,u,v,t,r)$	(x * t) / u
$f_i(x,u,v,t,r)$	r - x
$f_i(x,u,v,t,r)$	(v + t) / u
$f_i(x,u,v,t,r)$	ln(x + u + v + t+ r)

Para calcular el error cometido en la aproximación $f(x,u,v,t,r)$ en cada punto $(x_k, u_k, y_k, t_k, r_k)$ de la nube, se va a seguir el siguiente modelo:

$$E_k = f(x_k, u_k, v_k, t_k, r_k) - Y_k,$$

donde E_k es el error cometido en cada muestra, $f(x_k, u_k, v_k, t_k, r_k)$ es la aproximación obtenida, e Y_k es la salida esperada. Como se van a evaluar varias muestras, vamos a calcular el error medio cometido en la aproximación de la siguiente manera:

$$E_T = \frac{1}{s} * \sum_{k=1}^{k=s} E_k^2,$$

donde E_T es el error global cuadrático medio cometido en la aproximación, y s el número de muestras.

2.2.2 WxMaxima

Máxima es un sistema CAS (sistema de álgebra computacional), que puede resolver problemas matemáticos programando algoritmos y encontrando una fórmula que resuelva el problema, en lugar de generar solamente el valor numérico del resultado. Además de funcionar como una calculadora que representa variables numéricas, también ofrece varios métodos numéricos de análisis de ecuaciones, que no pueden resolverse analíticamente. Todo esto, presenta al usuario

mediante una interfaz gráfica, conocida como wxMáxima, esto permite a cualquier usuario exportar fácilmente tanto un libro de trabajo completo como el contenido de cualquier celda, como archivo de texto o HTML (WxMaxima, 2020).

Este programa nace en los años 70, como parte del proyecto MAC (Machine-Aided Cognition) en el MIT (Instituto de Tecnología de Massachusetts), y al principio se conoció como Macsymba. En los años 80, se entregó una copia al Departamento de Energía de Estados Unidos, y una versión del programa se distribuyó a las universidades, además, una copia fue mantenida por Bill Schelter, que se encargó de comercializarla, hasta el año de su muerte, que fue cuando se incluyó el proyecto en Sourceforge, donde se incluyeron nuevos participantes, código nuevo, y se aumentó el número de descargas (WxMaxima, 2020).

2.3 Redes Neuronales

Las redes neuronales artificiales se “construyen” normalmente con unidades informáticas idealizadas, que tienen valores de activación continuos y un conjunto de entradas ponderadas, estas unidades se conocen como neuronas, y su esquema básico se puede ver en la **figura 2a**. Estas neuronas pueden utilizar funciones de transferencia no lineales (funciones de salida), esto hace que sea más representativo apilar más de una capa de neuronas y hace posible el uso de métodos de optimización efectivos (ejecutables en tiempo real) para el entrenamiento (Tavanaei *et al.*, 2019).

La arquitectura básica de una red neuronal consiste en tres tipos de neuronas, neuronas de entrada, ocultas, y de salida, como se puede ver en la **figura 2b**. En las redes de avance, el flujo de señal sobre las unidades, es de entrada a salida, sólo en una dirección de avance, además, el procesamiento de datos puede extenderse sobre múltiples capas de unidades. Una red neuronal debe configurarse de modo que el conjunto de entradas, produzca el conjunto de salidas deseado. Una forma de configurar la red es establecer los pesos explícitamente, utilizando el conocimiento a priori. Otra forma es entrenar la red neuronal con patrones de enseñanza y dejando que cambie sus pesos de acuerdo con alguna regla de aprendizaje, dentro de esto, los tipos de aprendizaje se pueden dividir en supervisado, no supervisado y aprendizaje de refuerzo (Ajith Abraham, 1996).

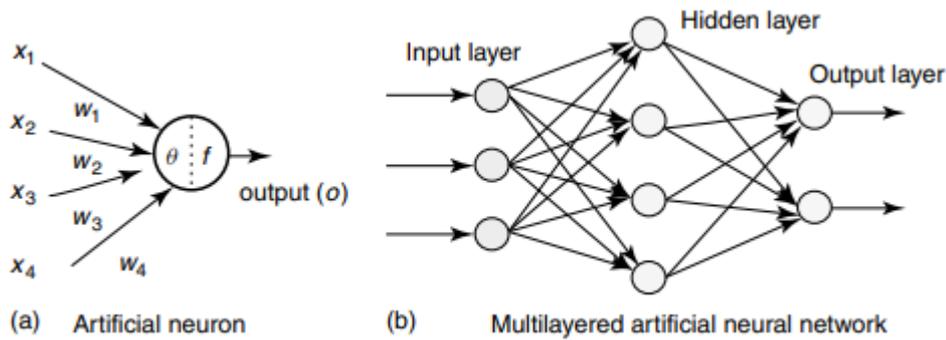


Figura 2. Esquema de una neurona artificial (a) y de una red neuronal artificial, según el algoritmo perceptrón multicapa (b) (Ajith Abraham, 1996).

2.3.1 Perceptrón Simple

El perceptrón simple es una red neuronal de una sola neurona no básica, bastante adecuada para resolver problemas simples en la clasificación de patrones. Supongamos que tenemos un conjunto de muestras de aprendizaje que consisten en un vector de entrada x y una salida deseada dk . Para una tarea de clasificación, la salida habitualmente es $+1$ o -1 . La regla de aprendizaje de perceptrón es muy simple y se establece de la siguiente manera:

1. Se empieza con pesos aleatorios.
2. Se selecciona un vector de entrada x del conjunto de entrenamiento.
3. Si la salida $yk \neq dk$ (el perceptrón da una respuesta incorrecta), por lo que se modifican los pesos w_i de acuerdo con:

$$\delta w_i = \eta (dk - yk) x_i; (\eta = \text{razón de aprendizaje}).$$
4. Se vuelve al paso 2, siempre y cuando no se cumpla el criterio de parada de la red.

Figura 3. Esquema del funcionamiento del algoritmo del perceptrón simple (Ajith Abraham, 1996).

2.3.2 Perceptrón Multicapa

Un perceptrón multicapa es una variante del perceptrón simple, pero que contiene una o más capas ocultas entre sus capas de entrada y salida. Las conexiones entre neuronas siempre se dirigen desde las capas inferiores a las superiores, y las neuronas en la misma capa no están interconectadas. El número de neuronas en la capa de entrada es igual al número de atributos de entrada, y el número de neuronas en la capa de salida es igual al número de clases. El proceso

de aprendizaje se basa en adaptar los pesos de las conexiones para obtener una diferencia mínima entre la salida obtenida y la esperada, y para esto, la técnica más utilizada es la retropropagación (Hassan Ramchoun, 2016).

2.3.3 Adaline

El algoritmo ADALINE (Adaptative Linear Neuron), es un método desarrollado en los años 60 por Bernar Widrow y Ted Hoff en la Universidad de Stanford. Este método tiene una estructura similar al perceptrón simple, pero se diferencian en la regla de aprendizaje que utilizan. El perceptrón usa como salida, la función umbral binaria, es decir, se tiene en cuenta solamente si se ha equivocado o no. Mientras que en Adaline, se utiliza como salida el número real obtenido por la red, por lo que se tiene en cuenta cuánto se ha equivocado, se va a utilizar la diferencia entre el valor real esperado y el producido por la red (Galván y Valls, 2019). El proceso de aprendizaje del algoritmo Adaline se puede ver en la **figura 4**.

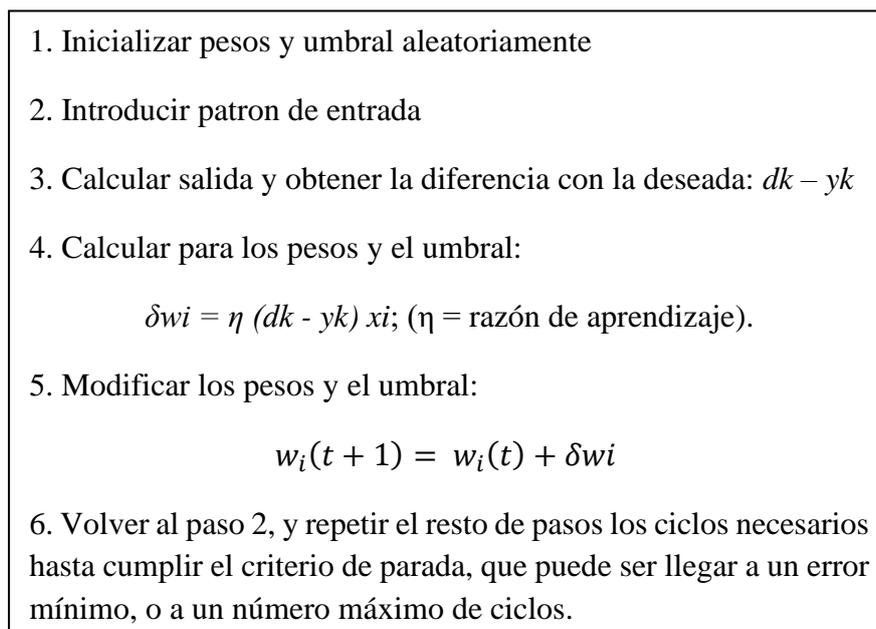


Figura 4. Esquema del funcionamiento del algoritmo Adaline (Galván y Valls, 2019).

2.3.4 Neuroph Studio

Neuroph es un software de redes neuronales escrito en lenguaje Java, con el que se pueden desarrollar distintas estructuras de redes neuronales. Neuroph simplifica el desarrollo de redes neuronales ya que contiene una biblioteca de redes neuronales Java y una herramienta GUI (interfaz gráfica de usuario) que permite la creación, capacitación y almacenamiento de redes neuronales (Java Neural Network Framework Neuroph, 2020).

En la **figura 5** se pueden encontrar las ventajas del uso de este software para el desarrollo de redes neuronales.

- | | |
|--|--|
| 1. Pequeño número de clases básicas esenciales en el paquete principal (solo 10) que pueden reutilizarse o ampliarse fácilmente. | 5. Soporte de reconocimiento de imagen. |
| 2. Soporte para reglas de aprendizaje supervisadas y no supervisadas. | 6. Soporte OCR. |
| 3. Una estructura y lógica fáciles de seguir. | 7. Muestra de predicción del mercado de valores. |
| 4. IDE de red neuronal y Java, <u>Neuroph Studio</u> , basado en la plataforma NetBeans. | 8. Muestras de visualización de aprendizaje. |
| | 9. Normalización de datos. |
| | 10. Marco simple de <u>microbenchmarking</u> |

Figura 5. Ventajas del uso de Neuroph Studio para la realización de pruebas mediante redes neuronales (Java Neural Network Framework Neuroph, 2020).

El esquema de funcionamiento de Neuroph Studio se puede ver en la **figura 6**, en primer lugar, se crea la red neuronal, seleccionando el número de neuronas de cada capa, la razón de aprendizaje, y la función de transferencia de las neuronas entre otros parámetros. Una vez configurada la red, se pasa al proceso de entrenamiento, para el cual es necesario introducir el conjunto de datos de entrenamiento. Cuando se entrena la red, obtenemos el valor del error del entrenamiento, si es demasiado alto, se modifican los parámetros necesarios de la red. Una vez entrenada la red, se pasa al paso de la validación, en el que se valida la configuración de la red mediante el conjunto de datos de validación.

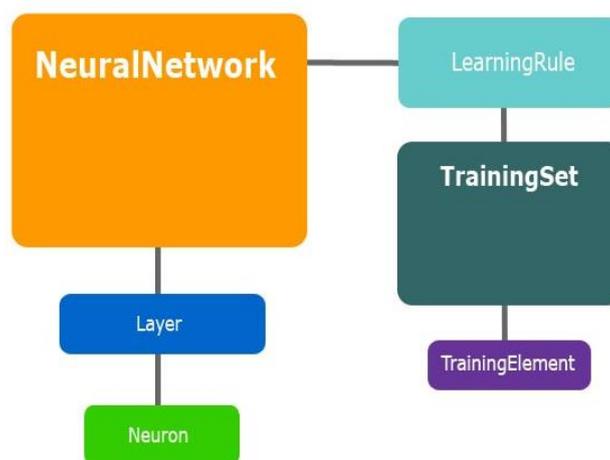


Figura 6. Esquema de funcionamiento del software Neuroph Studio (Java Neural Network Framework Neuroph, 2020).

3. RESULTADOS

A continuación, se van a presentar los valores obtenidos tras la realización de los métodos expuestos en el apartado anterior. Para cada prueba, se han obtenido diferentes parámetros que se han recogido y expuesto en distintas tablas y gráficos, los principales valores que se han utilizado, han sido los errores de entrenamiento y de validación del método, y la sensibilidad y especificidad de cada prueba, que son la relación entre los verdaderos positivos y los positivos totales, y los verdaderos negativos y los negativos totales respectivamente.

En la **tabla 3** se pueden ver el tipo de funciones utilizadas para cada prueba, que al no ser funciones lineales, porque los coeficientes a_i no multiplican directamente a las f_i , se tuvo que realizar la transformación en la salida para poder calcular la salida adecuadamente.

Tabla 3. Tipo de funciones utilizadas en cada prueba del método de regresión múltiple.

Prueba	Tipo de función
P1	$\frac{1}{1 + e^{-(a_0 * 1 + a_1 * \text{sen}(x+u) + a_2 * (v+t+r))}}$
P2	$\sqrt{\ln(a_0 * 1 + a_1 * \text{sen}(x+u) + a_2 * (v+t+r))} + 1$
P3	$\frac{1}{e^{(a_0 * 1 + a_1 * \text{sen}(x+u) + a_2 * (v+t+r) + a_3 * (e^{v-x} + r) + a_4 * (\frac{1}{u} + r))}}$
P4	$a_0 * 1 + a_1 * \text{sen}(x+u) + a_2 * (v+t+r) + a_3 * (e^{v-x} + r) + a_4 * \left(\frac{1}{u} + r\right)$
P5	$a_0 * 1 + a_1 * \text{sen}(x+u) + a_2 * (v+t+r) + a_3 * (e^{v-x} + r) + a_4 * \left(\frac{1}{u} + r\right) + a_5 * \frac{v}{r} + a_6 * \frac{x * t}{u}$
P6	$\frac{1}{1 + e^{-(a_0 * 1 + a_1 * \text{sen}(x+u) + a_2 * (v+t+r) + a_3 * (e^{v-x} + r) + a_4 * (\frac{1}{u} + r) + a_5 * \frac{v}{r} + a_6 * \frac{x * t}{u})}}$

Los resultados obtenidos mediante el método de regresión lineal se muestran en la **tabla 4**, como se puede ver, se ha conseguido tener una especificidad de método bastante alta, lo que indica que se detecta bastante bien a los individuos no enfermos, y respecto a la sensibilidad, al obtenerse que es de 0,8, se puede considerar que se detecta la enfermedad en la mayoría de los casos.

Tabla 4. Resultados de las pruebas realizadas usando el método de regresión múltiple, utilizando los conjuntos de entrenamiento y validación de la base de datos. Se ha utilizado como umbral del método 0,5.

Prueba	Error de entrenamiento	Error de validación	Sensibilidad	Especificidad
P1	0,181	0,154	0,6	0,8
P2	0,216	0,344	1	0
P3	0,188	0,086	0,6	1
P4	0,094	0,255	0,8	1
P5	0,091	0,305	0,8	1
P6	0,089	0,468	0,8	1

Como se puede ver en la tabla, se resalta la **prueba 3**, ya que es en la que se consiguieron unos resultados más equilibrados, aun así, se puede considerar que el error de entrenamiento del método es ligeramente alto.

Respecto a los errores cometidos en el método, se ha conseguido un error de entrenamiento por debajo del 0,1. Sin embargo, el error de validación es bastante alto. Por lo que se ha visto, los resultados obtenidos mediante este método son bastante buenos, pero hay varios inconvenientes de la regresión lineal múltiple, en primer lugar, en ocasiones hemos utilizado relaciones no lineales, por lo que se ha tenido que transformar la salida, además, se ha visto una limitación, ya que como wxMaxima no nos permite trabajar con un conjunto de tantos datos, se ha tenido que reducir bastante, tanto el conjunto de entrenamiento como el de validación, por lo que estas pruebas no son del todo representativas, aunque sí que nos han servido para tener una primera aproximación.

Debido a que se estuvieron manejando una gran cantidad de datos, se decidió utilizar el método de las redes neuronales para el análisis de la base de datos de imágenes mamográficas. Como se puede ver en la **tabla 5**, los resultados obtenidos con el método del perceptrón simple, mostraron un error de validación ligeramente alto, aunque presentaba tanto una buena especificidad, como una sensibilidad aceptable.

Tabla 5. Resumen de las pruebas realizadas para las redes neuronales, con los algoritmos Adaline y perceptrón simple. Se ha utilizado como umbral del método 0,5.

Tipo de red	Función de transferencia	Razón de aprendizaje	Error de entrenamiento	Error de validación	Sensibilidad	Especificidad
Perceptrón Simple	Step	0,2	0,086	0,212	0,72	0,84
Adaline	Lineal	0,2	0,23	0,235	0,47	0,97

Respecto a los resultados obtenidos mediante la red Adaline, que también se encuentran en la **tabla 5**, nos mostró un error de validación parecido al obtenido con el perceptrón simple, pero la diferencia se encontró en la sensibilidad y especificidad del método, ya que se vio que presentaba una especificidad cercana al valor máximo, que es 1, pero la sensibilidad no llegaba al 0,5; lo que indica que se detectó un número elevado de falsos negativos con este método.

Para las pruebas realizadas con el algoritmo perceptrón multicapa, se utilizaron varios parámetros con distintos valores, además de las distintas funciones de transferencia empleadas, también se utilizaron distintos valores de neuronas en la capa oculta, o distintas razones de aprendizaje en la red neuronal. En el **anexo 1** se puede ver la tabla de resultados mostrada por el software utilizado, Neuroph Studio, para la función sigmoide, con 25 neuronas en la capa oculta y 0,2 de razón de aprendizaje, en la tabla se aprecian los datos de entrada, la salida de la red, la salida esperada, y el error cometido.

En la **tabla 6**, se encuentra resumida la comparación de los resultados que se obtuvieron en las distintas pruebas realizadas con el perceptrón multicapa. En general se obtuvieron errores de entrenamiento cercanos al 0,1; y errores de validación ligeramente mayores a los de entrenamiento. Respecto a la sensibilidad y especificidad del método, se buscó la combinación de parámetros para los que tanto la sensibilidad como la especificidad fuesen buenas.

Obsérvese que los mejores resultados se obtuvieron en las pruebas destacadas en la **tabla 6**, que corresponden con la función de transferencia lineal, 10 neuronas en la capa oculta y 0,2 como razón de aprendizaje, y la función sigmoide con 25 neuronas en la capa oculta y 0,2 de razón de aprendizaje.

Tabla 6. Comparación de los valores obtenidos mediante la red neuronal, utilizando el algoritmo de perceptrón multicapa, además de distintas funciones de transferencia. Se ha utilizado como umbral del método 0,5.

Función de transferencia	Neuronas capa oculta	Razón de aprendizaje	Error de Entrenamiento	Error de validación	Sensibilidad	Especificidad
Sigmoide	4	0,2	0,05	0,12	0,65	0,92
Step	4	0,2	0,1	0,18	0,94	0,71
Lineal	4	0,2	0,06	0,134	0,56	0,97
Lineal	10	0,2	0,06	0,13	0,62	0,95
Sigmoide	10	0,2	0,05	0,129	0,68	0,88
Sigmoide	10	0,5	0,05	0,153	0,66	0,88
Sigmoide	10	0,5	0,05	0,144	0,65	0,88
Sigmoide	10	0,5	0,05	0,13	0,71	0,87
Step	10	0,2	0,103	0,27	0,93	0,57
Sigmoide	25	0,2	0,053	0,126	0,72	0,88
Sigmoide	25	0,5	0,051	0,131	0,69	0,88
Step	25	0,2	0,106	0,247	0,47	1
Step	25	0,5	0,107	0,219	0,53	1
Lineal	25	0,2	0,056	0,235	0,63	0,88
Step	4	0,5	0,12	0,226	0,6	0,922

Para el análisis ROC del método, se eligieron las dos mejores combinaciones de parámetros, resaltados en la **tabla 6**, en primer lugar, se utilizó la función lineal como función de transferencia, utilizando 10 neuronas en la capa oculta y una razón de aprendizaje de 0,2. En segundo lugar, se escogió la función sigmoide, 25 neuronas en la capa oculta y 0,2 como razón de aprendizaje.

Tabla 7. Resultados del análisis ROC para las dos pruebas escogidas.

Función de transferencia	Sensibilidad	Especificidad	Ecuación de la curva	A(z)
Lineal (umbral 0,3)	0,88	0,79	$y = -2,7381x^2 + 2,3369x + 0,51$	0,766
Lineal (umbral 0,5)	0,62	0,95		
Lineal (umbral 0,7)	0,51	1		
Sigmoide (umbral 0,3)	0,93	0,71	$y = -2,5588x^2 + 2,2844x + 0,4827$	0,772
Sigmoide (umbral 0,5)	0,72	0,88		
Sigmoide (umbral 0,7)	0,57	0,96		

En la **tabla 7**, se muestran los resultados que se obtuvieron con el análisis ROC, para las pruebas seleccionadas con anterioridad. En la tabla se muestra la ecuación obtenida para cada prueba, que se pueden ver en las **figuras 7 y 8**, además, se calculó el valor $A(z)$, que se trata del área bajo la curva, que es en ambos casos cercano al 0,77; lo que nos indica que el método tiene una probabilidad del 77% de clasificar bien las muestras.

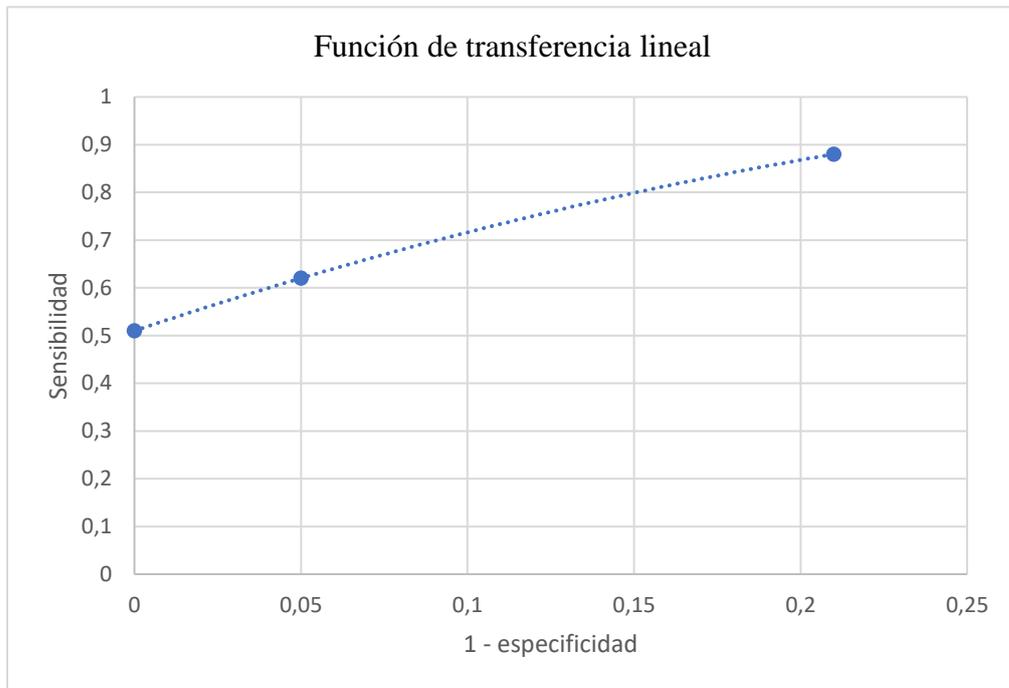


Figura 7. Representación gráfica del valor 1 – especificidad frente a la sensibilidad para el método, utilizando la función de transferencia lineal, y tomando como valores umbrales del método 0’3, 0’5 y 0’7.

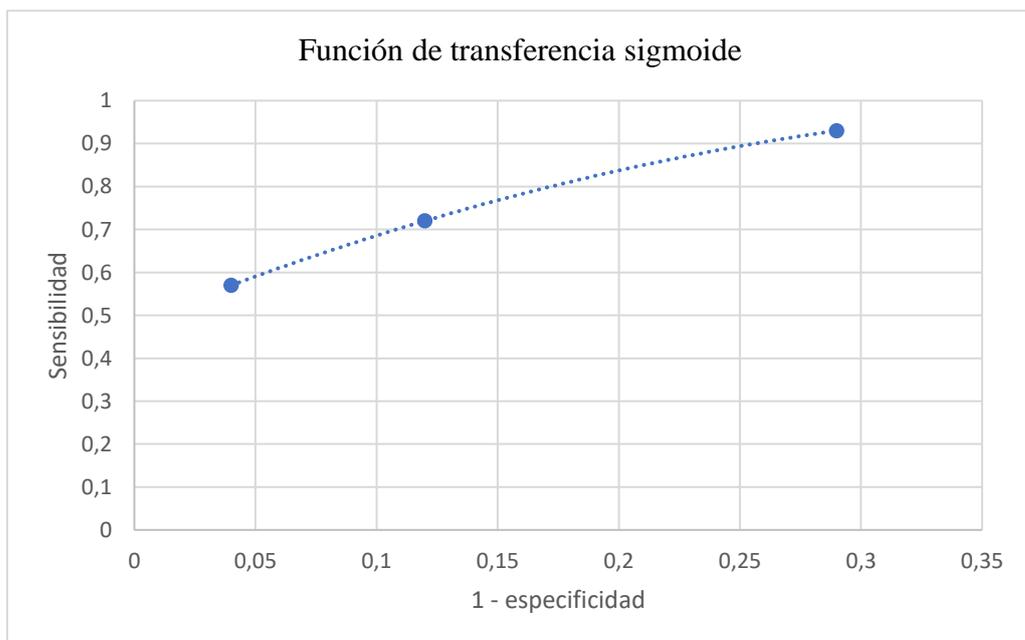


Figura 8. Representación gráfica del valor 1 – especificidad frente a la sensibilidad para el método, utilizando la función de transferencia sigmoide, y tomando como valores umbrales del método 0’3, 0’5 y 0’7.

4. CONCLUSIONES

En este trabajo se han presentado dos sistemas para lograr una predicción de los resultados de biopsias de cáncer de mama, a partir de la imagen de la mamografía realizada a la paciente, y la introducción de una serie de atributos como la edad de la paciente, el parámetro BI-RADS, o diferentes datos sobre la forma y el margen de la masa detectada en la mamografía.

Los dos sistemas utilizados han sido la regresión lineal múltiple y redes neuronales. En la regresión lineal, probando con distintos tipos de funciones, y en las redes neuronales con el uso de diferentes tipos de algoritmos, perceptrón simple, Adaline y perceptrón multicapa, y cambiando los distintos parámetros para la red. Se han evaluado ambos sistemas mediante la sensibilidad y especificidad para cada método, y en el caso de las redes neuronales, también se ha utilizado el análisis ROC.

En primer lugar, el método de regresión lineal múltiple ha presentado una buena sensibilidad y especificidad, lo que nos hace suponer que puede favorecer la detección de la enfermedad, pero presenta un inconveniente, y es que el número de casos utilizados para el análisis de este método es relativamente pequeño, debido a las limitaciones que presenta tanto el método como el software utilizado para el desarrollo de este. Por lo tanto, se necesitarían de más investigaciones, y la obtención de más resultados para poder asegurar la fiabilidad y viabilidad de este método para la detección de masas carcinógenas malignas, sin la necesidad de realizar una biopsia a la paciente.

El segundo método utilizado en este trabajo, han sido las redes neuronales artificiales, y dentro de este, se han empleado los algoritmos perceptrón simple, Adaline, y perceptrón multicapa. Tras el análisis de los resultados obtenidos para los tres enfoques de la red neuronal artificial, se obtuvieron unos mejores resultados para el perceptrón multicapa.

Debido a esto, se realizaron distintos enfoques para la red perceptrón multicapa, modificando las funciones de transferencia de las neuronas, y además, dentro de cada prueba también se modificaron distintos parámetros de la red como por ejemplo, la razón de neuronas, o el número de neuronas de la capa oculta de la red.

Para cada prueba del algoritmo del perceptrón multicapa, se realizó un análisis de la efectividad del método, mediante el cálculo del error que cometía la red al proporcionar una respuesta, y además, un análisis de la especificidad y la sensibilidad de cada enfoque específico del método.

Los resultados nos mostraron que los mejores modelos para el perceptrón multicapa, eran utilizando las funciones de transferencia lineal y sigmoide. Para el análisis de los resultados de ambos métodos, se realizó un análisis ROC, en el que se calculó el valor $A(z)$, que es el área bajo la curva del gráfico en el que se representa la sensibilidad frente al valor de $1 - \text{especificidad}$ de ambos métodos, como se representa en las **figuras 7 y 8**. Ambos enfoques presentaron buenos valores de predicción, llegando al 77% de fiabilidad del método.

El conjunto de estos resultados nos hace esperar que el algoritmo del perceptrón multicapa es el mejor enfoque para la reducción del número de biopsias en pacientes de cáncer de mama, sin embargo, los resultados de este trabajo se han obtenido con un número de pacientes relativamente bajo, cercano a los mil casos, por lo que se necesitarían realizar más estudios y análisis con una base de datos más grande, de esta manera se podría llegar a obtener un método de detección lo demasiado confiable como para no tener que realizar biopsias, y detectar tumores malignos sin realizar procesos quirúrgicos tan invasivos.

5. REFERENCIAS

- Ajith Abraham (1996) "Artificial Neural Networks ", *Journal of biomedicine biotechnology*. Chichester, UK: John Wiley & Sons, Ltd, 91(S8), pp. 1615–1635. doi:10.1002/0471497398.mm421.
- Akram, M., Iqbal, M., Daniyal, M. and Khan, A. U. (2017) "Awareness and current knowledge of breast cancer", *Biological Research*. BioMed Central Ltd. doi:10.1186/s40659-017-0140-9.
- Aronowitz, R. A. (2007) "Unnatural History: Breast Cancer and American Society", *Cambridge University Press*, p. 366.
- Balleyguier, C., Ayadi, S., Van Nguyen, K., Vanel, D., Dromain, C. and Sigal, R. (2007) "BIRADS™ classification in mammography", *European Journal of Radiology*, 61(2), pp. 192–194. doi:10.1016/j.ejrad.2006.08.033.
- Bangdiwala, S. I. (2018) "Regression: multiple linear", *International Journal of Injury Control and Safety Promotion*. Taylor and Francis Ltd., 25(2), pp. 232–236. doi:10.1080/17457300.2018.1452336.
- Dua, D. y Graff, C. (2019). *UCI Machine Learning Repository*. Disponible en: <http://archive.ics.uci.edu/ml> (Accedido: 20 de febrero de 2020)
- Elter, M., Schulz-Wendtland, R. and Wittenberg, T. (2007) "The prediction of breast cancer biopsy outcomes using two CAD approaches that both emphasize an intelligible decision process", *Medical Physics*, 34(11), pp. 4164–4172. doi:10.1118/1.2786864.

Galván, I.M. y Valls J.M. (2019) *Primeros Modelos Computacionales*. Curso redes neuronales artificiales. Universidad Carlos III de Madrid.

Hassan Ramchoun, M. A. J. I. Y. G. M. E. (2016) "Multilayer Perceptron: Architecture Optimization and Training", *International Journal of Interactive Multimedia and Artificial Intelligence*, 4, pp. 26–30. doi:10.9781/ijimai.2016.415.

Java Neural Network Framework Neuroph (2020) [*Página web de los desarrolladores*]. Disponible en: <http://neuroph.sourceforge.net/> (Accedido: 24 de enero de 2020).

Kriegeskorte, N. y Golan, T. (2019) "Neural network models and deep learning", *Current Biology*. Cell Press, 29(7), pp. R231–R236. doi:10.1016/j.cub.2019.02.034.

Løberg, M., Lousdal, M. L., Bretthauer, M. and Kalager, M. (2015) "Benefits and harms of mammography screening", *Breast Cancer Research*. BioMed Central Ltd. doi:10.1186/s13058-015-0525-z.

Merino Bonilla, J. A., Torres Tabanera, M. and Ros Mendoza, L. H. (2017) "Breast cancer in the 21st century: from early detection to new therapies", *Radiología*. Ediciones Doyma, S.L., 59(5), pp. 368–379. doi:10.1016/j.rx.2017.06.003.

Reig, B., Heacock, L., Geras, K. J. and Moy, L. (2019) "Machine learning in breast MRI", *Journal of Magnetic Resonance Imaging*. John Wiley and Sons Inc. doi:10.1002/jmri.26852.

Tavanaei, A., Ghodrati, M., Kheradpisheh, S. R., Masquelier, T. and Maida, A. (2019) "Deep learning in spiking neural networks", *Neural Networks*. Elsevier Ltd, 111, pp. 47–63. doi:10.1016/j.neunet.2018.12.002.

Universidad de Belgrado (2020) *Neuroph Studio (Versión 2.98)* [*Programa de ordenador*]. Disponible en: <http://neuroph.sourceforge.net/download.html> (Accedido: 15 de febrero de 2020).

WxMaxima (2020) [*Página web de los desarrolladores del software*]. Disponible en: <https://wxmaxima-developers.github.io/wxmaxima/index.html> (Accedido: 16 de marzo de 2020).